

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of )  
Yoko FUJIWARA ) Group Art Unit: Unassigned  
Application No.: Unassigned ) Examiner: Unassigned  
Filed: June 14, 2001 )  
For: IMAGE PROCESSING DEVICE AND )  
PROGRAM PRODUCT )  
)  
)  
)  
)



CLAIM FOR CONVENTION PRIORITY

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

The benefit of the filing date of the following prior foreign application in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed:

Japanese Patent Application No. 2000-185267

Filed: June 20, 2000

In support of this claim, enclosed is a certified copy of said prior foreign application. Said prior foreign application was referred to in the oath or declaration. Acknowledgment of receipt of the certified copy is requested.

Respectfully submitted,

BURNS, DOANE, SWECKER & MATHIS, L.L.P.

Date: June 14, 2001

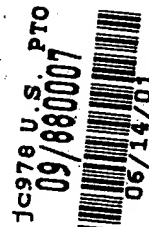
By:

Platon N. Mandros  
Registration No. 22,124

P.O. Box 1404  
Alexandria, Virginia 22313-1404  
(703) 836-6620

日 本 国 特 許 庁

PATENT OFFICE  
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2000年 6月20日

出 願 番 号

Application Number:

特願2000-185267

出 願 人

Applicant (s):

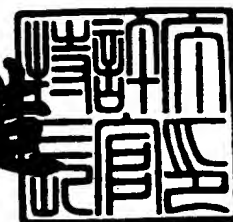
ミノルタ株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 3月 2日

特許庁長官  
Commissioner,  
Patent Office

及 川 耕 造



出証番号 出証特2001-3015574

【書類名】 特許願

【整理番号】 AK05219

【提出日】 平成12年 6月20日

【あて先】 特許庁長官 近藤 隆彦 殿

【国際特許分類】 G06K 9/20

【発明の名称】 画像認識装置、画像認識方法および画像認識プログラム  
を記録したコンピュータ読取可能な記録媒体

【請求項の数】 11

【発明者】

    【住所又は居所】 大阪府中央区安土町二丁目3番13号 大阪国際ビル  
                        ミノルタ株式会社内

    【氏名】 藤原 葉子

【特許出願人】

    【識別番号】 000006079

    【氏名又は名称】 ミノルタ株式会社

【代理人】

    【識別番号】 100072349

    【弁理士】

    【氏名又は名称】 八田 幹雄

    【電話番号】 03-3230-4766

【選任した代理人】

    【識別番号】 100102912

    【弁理士】

    【氏名又は名称】 野上 敦

【選任した代理人】

    【識別番号】 100110995

    【弁理士】

    【氏名又は名称】 奈良 泰男

【選任した代理人】

【識別番号】 100111464

【弁理士】

【氏名又は名称】 齋藤 悦子

【選任した代理人】

【識別番号】 100114649

【弁理士】

【氏名又は名称】 宇谷 勝幸

【手数料の表示】

【予納台帳番号】 001719

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像認識装置、画像認識方法および画像認識プログラム  
を記録したコンピュータ読取可能な記録媒体

【特許請求の範囲】

【請求項 1】 画像認識装置において、

画像データ中の文字画像から文字コードを認識する文字認識部と、

文字コードに応じて、文字画像を文字コードデータに変換する変換部と、

文字認識部における認識の確からしさである文字認識確度を検出する文字認識  
確度検出部と、

個々の文字画像が複数集合して形成された文字画像集合の中に、文字認識確度  
が所定確度より低い文字画像が含まれるか否かを検出する低確度文字画像検出部  
と、

前記文字画像集合の中に低確度文字画像が含まれている場合、当該文字画像集  
合全体に対して文字コードデータへの変換を制止する制御部と、

を有することを特徴とする画像認識装置。

【請求項 2】 画像認識装置において、

画像データ中の文字画像から文字コードを認識する文字認識部と、

文字コードに応じて、文字画像を文字コードデータに変換する変換部と、

文字認識部における認識の確からしさである文字認識確度を検出する文字認識  
確度検出部と、

文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る  
並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在するか否  
かを検出する特定文字画像検出部と、

前記特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が  
複数集合して形成された文字画像集合全体に対して文字コードデータへの変換を  
制止する制御部と、

を有することを特徴とする画像認識装置。

【請求項 3】 前記特定文字画像は、斜体を呈し、前記文字画像集合は、当  
該特定文字画像を含む単語を形成していることを特徴とする請求項 2 に記載の画

像認識装置。

【請求項 4】 前記特定文字画像は、隣接する文字画像とカーニングが行われる組を形成し、前記文字画像集合は、少なくとも当該特定文字画像とカーニングが行われる隣接する文字画像との組を含むことを特徴とする請求項 2 に記載の画像認識装置。

【請求項 5】 電子ファイルを作成するファイル作成部、をさらに有することを特徴とする請求項 1 または 2 に記載の画像認識装置。

【請求項 6】 原稿を読み取ることによって画像データを得るスキャナ部、をさらに有することを特徴とする請求項 1 または 2 に記載の画像認識装置。

【請求項 7】 画像データを受信する画像データ受信部と、データを用紙に印刷するプリント部と、をさらに有することを特徴とする請求項 1 または 2 に記載の画像認識装置。

【請求項 8】 画像認識方法において、  
画像データ中の文字画像から文字コードを認識する過程と、  
文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、  
個々の文字画像が複数集合して形成された文字画像集合の中に、文字認識確度が所定確度より低い文字画像が含まれるか否かを検出する過程と、  
前記文字画像集合の中に低確度文字画像が含まれている場合、当該文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、  
を有することを特徴とする画像認識方法。

【請求項 9】 画像認識方法において、  
画像データ中の文字画像から文字コードを認識する過程と、  
文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、  
文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在するか否かを検出する過程と、

前記特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が複数集合して形成された文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、

を有することを特徴とする画像認識方法。

【請求項 1 0】 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、

画像データ中の文字画像から文字コードを認識する過程と、

文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、

個々の文字画像が複数集合して形成された文字画像集合の中に、文字認識確度が所定確度より低い文字画像が含まれるか否かを検出する過程と、

前記文字画像集合の中に低確度文字画像が含まれている場合、当該文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、

をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

【請求項 1 1】 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、

画像データ中の文字画像から文字コードを認識する過程と、

文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、

文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在するか否かを検出する過程と、

前記特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が複数集合して形成された文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、

をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

## 【発明の詳細な説明】

【0001】

## 【発明の属する技術分野】

本発明は、画像認識技術に関し、特に、文字コードを認識する確からしさの度合いが小さい文字画像を、文字コードデータに変換することなく文字画像から切り出された文字画像データで出力する画像認識技術に関する。

【0002】

## 【従来の技術】

最近の画像認識装置は、読取に関する条件の良い原稿（例えば同一フォントの文字のみで構成された原稿）であれば、極めて高い精度で文字画像を文字コードとして認識可能である。しかしながら、原稿上の文字の品質が悪かったり、原稿上の文字等のレイアウトが複雑であると、文字画像を文字コードとして認識する認識精度が大きく低下し、誤認識される文字画像が増加してしまう。

【0003】

このような問題に対し、誤認識されている可能性が高い文字画像を文字コードデータに変換することなく、文字画像データ（例えばビットマップ形式の画像データ）のまま出力する画像認識装置が提案されている。この画像認識装置によれば、原稿上の文字とは異なる文字コードデータを出力してしまうことがない。

【0004】

## 【発明が解決しようとする課題】

しかしながら、誤認識されている可能性が高い文字画像のみを文字画像データとして出力した場合、図12に示すように文字画像データと文字コードデータとの形状に差異が生じてしまい、ユーザに違和感を与えるといった問題があった（図中矩形で示す部分は、文字画像データとして出力した部分）。

【0005】

図13（A）は、カーニングが行われている文字画像を従来の画像認識装置で出力した場合を示しており、文字画像データと文字コードデータとの差異が顕著である。

【0006】



カーニングとは、隣り合う特定文字の組合わせで印字するときに、文字間の距離を詰めて空間のバランスの取れた文字配列をすることであり、図 1 3 (A) では文字画像データである「W」の右下部分に文字コードデータの「e」を詰めて配列した例を示している。文字コードデータ「e」の左側の一部分が文字画像データ「W」によって隠れてしまっている。

## 【 0 0 0 7 】

また、図 1 3 (B) は、斜体で表現されている文字画像を出力した場合を示しており、同様に差異が顕著である。

## 【 0 0 0 8 】

斜体で表現された「W」の文字画像データの右下側には文字画像「e」の左側の一部分が含まれており、文字コードデータで出力された「e」との位置がずれている。また、斜体で表現された「n」の文字画像データの左上部分に文字画像「k」の右側の一部分が含まれており、文字コードデータで出力された「k」との位置がずれている。

## 【 0 0 0 9 】

本発明の目的は、画像データ中の文字画像から文字コードを認識して得られた文字コードデータと文字画像から文字コードデータに変換することなく切り出された文字画像データとが混在する画像を出力する画像認識装置において、双方の混在によりユーザに与える違和感を低減することである。

## 【 0 0 1 0 】

## 【課題を解決するための手段】

本発明の目的は、下記する手段により達成される。

## 【 0 0 1 1 】

(1) 画像認識装置において、画像データ中の文字画像から文字コードを認識する文字認識部と、文字コードに応じて、文字画像を文字コードデータに変換する変換部と、文字認識部における認識の確からしさである文字認識確度を検出する文字認識確度検出部と、個々の文字画像が複数集合して形成された文字画像集合の中に、文字認識確度が所定確度より低い文字画像が含まれるか否かを検出する低確度文字画像検出部と、前記文字画像集合の中に低確度文字画像が含まれ

ている場合、当該文字画像集合全体に対して文字コードデータへの変換を制止する制御部と、を有することを特徴とする画像認識装置。

【 0 0 1 2 】

(2) 画像認識装置において、画像データ中の文字画像から文字コードを認識する文字認識部と、文字コードに応じて、文字画像を文字コードデータに変換する変換部と、文字認識部における認識の確からしさである文字認識確度を検出する文字認識確度検出部と、文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在するか否かを検出する特定文字画像検出部と、前記特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が複数集合して形成された文字画像集合全体に対して文字コードデータへの変換を制止する制御部と、を有することを特徴とする画像認識装置。

【 0 0 1 3 】

(3) 前記特定文字画像は、斜体を呈し、前記文字画像集合は、当該特定文字画像を含む単語を形成していることを特徴とする上記(2)に記載の画像認識装置。

【 0 0 1 4 】

(4) 前記特定文字画像は、隣接する文字画像とカーニングが行われる組を形成し、前記文字画像集合は、少なくとも当該特定文字画像とカーニングが行われる隣接する文字画像との組を含むことを特徴とする上記(2)に記載の画像認識装置。

【 0 0 1 5 】

(5) 電子ファイルを作成するファイル作成部、をさらに有することを特徴とする上記(1)または(2)に記載の画像認識装置。

【 0 0 1 6 】

(6) 原稿を読み取ることによって画像データを得るスキャナ部、をさらに有することを特徴とする上記(1)または(2)に記載の画像認識装置。

【 0 0 1 7 】

(7) 画像データを受信する画像データ受信部と、データを用紙に印刷する

プリント部と、をさらに有することを特徴とする上記（１）または（２）に記載の画像認識装置。

## 【 0 0 1 8 】

（８） 画像認識方法において、画像データ中の文字画像から文字コードを認識する過程と、文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、個々の文字画像が複数集合して形成された文字画像集合の中に、文字認識確度が所定確度より低い文字画像が含まれるか否かを検出する過程と、前記文字画像集合の中に低確度文字画像が含まれている場合、当該文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、を有することを特徴とする画像認識方法。

## 【 0 0 1 9 】

（９） 画像認識方法において、画像データ中の文字画像から文字コードを認識する過程と、文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在するか否かを検出する過程と、前記特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が複数集合して形成された文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、を有することを特徴とする画像認識方法。

## 【 0 0 2 0 】

（１０） 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、画像データ中の文字画像から文字コードを認識する過程と、文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、個々の文字画像が複数集合して形成された文字画像集合の中に、文字認識確度が所定確度より低い文字画像が含まれるか否かを検出する過程と、前記文字画像集合の中に低確度文字画像が含まれている場合、当該文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

## 【 0 0 2 1 】

(11) 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、画像データ中の文字画像から文字コードを認識する過程と、文字コードを認識する際の認識の確からしさである文字認識確度を検出する過程と、文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在するか否かを検出する過程と、前記特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が複数集合して形成された文字画像集合全体に対して文字コードに応じた文字コードデータへの変換を制止する過程と、をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

【0022】

【発明の実施の形態】

以下、添付した図面を参照して、本発明の実施形態を説明する。

【0023】

図1は、本発明の第1実施形態に係る画像認識装置を含む情報機器システムの構成図である。

【0024】

情報機器システム100は、画像認識装置としての機能を有するコンピュータ10、画像入力装置であるスキャナ50、および、画像出力装置であるプリンタ60等の情報機器がネットワークN上に接続されて構成されている。

【0025】

コンピュータ10、スキャナ50、および、プリンタ60は、ネットワークNを経由して互いにデータの授受を行うことができる。

【0026】

図2は、コンピュータ10の概略構成を示すブロック図である。

【0027】

コンピュータ10は、各情報機器との間でデータの授受を行うためのインターフェース(I/F)18と、スキャナ50を制御するためのスキャナドライバ12と、プリンタ60を制御するためのプリンタドライバ14と、所定のプログラムやデータを記憶するROM15と、一時的にデータを記憶するRAM16と、

各種情報の表示やユーザからの指示入力可能な操作パネル 1 7 と、文字認識等の処理を行う画像処理部 1 3（詳細後述）と、上記各部を制御する主制御部 1 1 とを有している。なお、スキャナドライバ 1 2、プリンタドライバ 1 4 は、ソフトウェアであり、図示しないハードディスク等の記憶装置に記憶されている。

#### 【 0 0 2 8 】

図 1 に示す情報機器システムの動作指示は、コンピュータ 1 0 の操作パネル 1 7 から行うことができる。本実施形態において操作パネル 1 7 は、コンピュータ 1 0 のディスプレイ上に表示される架空のパネルであり、コンピュータ 1 0 を操作するためのキーボードやマウス等の機器により、動作指示を行う。

#### 【 0 0 2 9 】

例えば、操作パネル 1 7 上にはスキャン&ファイルボタンとスキャン&プリントボタンとが表示されており、スキャン&ファイルボタンで動作指示を行った場合には、スキャナ 5 0 は原稿読取動作を開始し、読み取りによって得られた画像データはコンピュータ 1 0 に送信され、画像処理部 1 3 にて文字認識等の画像処理がなされた後、所定の形式のファイルとしてコンピュータ 1 0 内の記憶装置に保存される。

#### 【 0 0 3 0 】

また、スキャン&プリントボタンで動作指示を行った場合には、スキャナ 5 0 の読み取りによって得られた画像データは、画像処理部 1 3 にて文字認識等の画像処理がなされた後、プリンタ 6 0 に送信され用紙上に印刷される。

#### 【 0 0 3 1 】

次に、画像処理部 1 3 について詳細に説明する。

#### 【 0 0 3 2 】

画像処理部 1 3 は、文字認識部 3 1、文字色検出部 3 2、文字出力形式判定部 3 3、文字画像データ作成部 3 4、文字消去部 3 5、およびファイル作成部 3 6 を備えている。

#### 【 0 0 3 3 】

文字認識部 3 1 は、個々の文字画像から文字コードを認識して文字コードデータを得るほか、文字コードを認識する確からしさの度合いである文字認識確度を

検出する。文字画像から文字コードを認識する方法は、例えば、各文字画像の特徴量とあらかじめ記憶されている辞書パターンとの一致の度合いに基づいて行われる。文字認識確度は、例えば 10 段階で与えられ、値が大きいと文字認識確度が高いことを意味する。また、文字認識部 3 1 は、文字画像の位置情報を認識する。個々の文字画像の位置情報は、例えば図 3 に示すように、画像データの左上を原点として、各文字画像の外接矩形の左上および右下の座標として得られる。さらに、文字認識部 3 1 は、文字画像から、フォントの種類やフォントサイズ、さらには斜体などのスタイル、等の文字属性を認識する。フォントは、所定の数種類の代表的な出力フォントに分類されて判別される。また、文字色検出部 3 2 は、文字コードが認識された個々の文字画像の色を検出する。

## 【 0 0 3 4 】

文字出力形式判定部 3 3 は、画像データ中の文字画像を、文字コードデータに変換するか、あるいは文字コードデータに変換せずに、文字画像を切り出して文字画像データを作成するかの判定を行う。そして、文字画像データ作成部 3 4 は、文字画像データを作成する旨の判定が行われた文字画像を、入力された画像データから切り出して文字画像データを作成する。

## 【 0 0 3 5 】

文字消去部 3 5 は、元の画像データ、すなわち入力されたカラーの画像データから、文字コードデータまたは文字画像データとされた後の元の文字画像を消去する。

## 【 0 0 3 6 】

ファイル作成部 3 6 は、入力された画像データ中の文字画像から得られた文字コードデータや文字画像データを用いて、所定の形式のファイルを作成する。

## 【 0 0 3 7 】

次に、画像処理部 1 3 の動作を図 4 および図 5 のフローチャートを用いて説明する。

## 【 0 0 3 8 】

まず、スキャナ 5 0 で原稿を読み取ることによって得られた画像データに対し、RGB 表色系から Lab 表色系へと色空間の変換を行う (S 1 1)。Lab 表

色系は、スキャナあるいはプリンタ等の個々の情報機器に依存しないので、ネットワーク上で共通の色情報として扱うことができるものである。

#### 【 0 0 3 9 】

次に、文字認識部 3 1 に入力する画像データを作成するために、OCR 前処理が行われる (S 1 2)。後に文字認識部 3 1 がモノクロ 2 値画像データから特徴量を抽出して文字コードの認識を行うことから、OCR 前処理では、まず、L a b 表色系で表示されたカラー画像データに対して 2 値化処理を行うことにより、文字画像が黒で、その背景画像が白のモノクロ 2 値画像データを作成する。本実施形態では、文字画像およびその背景画像の L、a、b のレベルを検出することにより、例えば白地に書かれた色文字や、色地に書かれた黒文字等が原稿上に表されていても、原稿を読み取って得られた画像データ中の文字画像から十分に文字コードを認識することができるように 2 値化処理を行う。この他、OCR 前処理では、入力された画像データ中の孤立点等のノイズの除去、原稿が傾いて読み取られた場合の補正処理、つぶれたりかすれたりしている文字画像の補正なども行われる。

#### 【 0 0 4 0 】

OCR 前処理により得られたモノクロ 2 値画像データは、文字認識部 3 1 に入力され、文字認識処理が行われる (S 1 3)。文字認識部 3 1 は、モノクロ 2 値画像データから文字画像が存在する文字領域の抽出を行った後、文字領域内の個々の文字画像から文字コードを認識する。また、文字認識部 3 1 は、文字コードを認識する確からしさの度合いである文字認識確度を検出することができ、さらに、文字画像の位置情報のほか、フォントの種類、フォントサイズ、斜体などのスタイル等の文字属性を認識する。この結果、文字認識部 3 1 からは、個々の文字画像に対して、文字コードデータその他、文字認識確度、位置情報、およびフォントの種類等の文字属性などの情報が出力される。

#### 【 0 0 4 1 】

次いで、文字色検出処理が行われる (S 1 4)。すなわち、文字コードが認識された個々の文字画像の色は、文字色検出部 3 2 により検出される。ここでは、L a b 表色系に色変換された画像データが使用される。文字画像の色の検出は、

例えば、画像データ中の文字画像が存在する文字領域の L a b の値を読み取り、文字画像を構成する全画素について、L、a、b それぞれの平均値を求めることにより行う。文字画像とその背景画像との画素の区別は、OCR 前処理 (S 1 2) で得られたモノクロ 2 値画像データを用いて行う。つまり、個々の文字画像の外接矩形内において、モノクロ 2 値画像データにおける黒画素を文字画像の画素、白画素を背景画像の画素と区別することができる。

#### 【0042】

文字画像を文字コードデータに変換して出力する場合、文字色検出部 3 2 により検出された個々の文字画像の色情報に基づいて、色を指定して文字コードデータが出力される。これにより、例えば原稿上の文字の色を再現することができる。

#### 【0043】

そして、画像データ中の個々の文字画像について出力形式の判定処理が行われる (S 1 5)。すなわち、文字出力形式判定部 3 3 により、画像データ中の文字画像を、文字コードデータに変換するか、あるいは文字コードデータに変換せずに、文字画像を切り出して文字画像データを作成するかの判定が行われる。

#### 【0044】

文字出力形式判定部 3 3 は、まず、文字認識確度が所定値以上の文字画像を、文字コードデータに変換する候補の文字画像とし、文字認識確度が所定値より小さい文字画像を、文字画像データを作成する候補の文字画像とする。そして、文字出力形式判定部 3 3 は、単語や文字画像ブロック内での後述する特別な文字出力形式の判定処理が行われない場合には、個々の文字画像ごとに、候補とされた内容通りの文字出力形式の判定を行う。すなわち、文字認識確度が所定値以上の文字コードデータ候補の文字画像は、文字コードデータで出力し、文字認識確度が所定値より小さい文字画像データ候補の文字画像は、文字画像データで出力する旨の文字出力形式の判定が行われる。

#### 【0045】

本実施形態の文字出力形式判定部 3 3 は、個々の文字画像が複数集合して形成された文字画像集合として、単語と、パラグラフ程度に文字画像が存在する集合



である文字画像ブロックとを取り上げ、それらの文字画像集合の中に、文字認識確度が所定値より小さい文字画像データ候補の文字画像がどの程度混在しているかを検出して、以下に説明するように文字コードデータと文字画像データとを違和感なく出力させるための判定を行う。

## 【 0 0 4 6 】

具体的には、図 5 に示すように、まず、注目する単語の順番を示す変数  $j$  と、注目する文字画像ブロックの順番を示す変数  $k$  とを初期化する (S 2 1)。

## 【 0 0 4 7 】

次に、注目単語についての出力形式の判定を開始する。すなわち、注目単語内の個々の文字画像の中で、例えば 1 0 段階で表される文字認識確度  $C$  が所定の閾値  $Th1$  より小さい文字画像データ候補の文字画像がいくつあるかをカウントし、 $Nb$  とする (S 2 2)。なお、閾値  $Th1$  は、適宜設定することができる。

## 【 0 0 4 8 】

注目単語内において、文字画像データ候補の文字画像の数  $Nb$  が所定の閾値  $Th2$  より大きい場合 (S 2 3 で Y E S)、注目単語内の全文字画像を切り出して文字画像データを作成する旨の出力形式の判定を行う (S 2 4)。なお、閾値  $Th2$  は、適宜設定することができ、例えば文字画像データ候補の文字画像の数が 1 個以上あった場合でも注目単語内の全文字画像を切り出して文字画像データを作成すると判定させることが可能である。また、閾値  $Th2$  は、例えば単語内の文字画像の総数によって変化させることも可能である。さらに、注目単語内における文字画像データ候補の文字画像の数  $Nb$  を判定に使用する代わりに、数  $Nb$  を単語内の文字画像の総数で除した割合を判定に使用することも勿論可能である。

## 【 0 0 4 9 】

そして、注目する単語の順番を示す変数  $j$  を繰り上げ (S 2 5)、画像データ内の全単語について、上記の処理を繰り返し行う (S 2 6)。

## 【 0 0 5 0 】

次に、注目文字画像ブロックについての出力形式の判定を開始する。本実施形態では、文字画像ブロック内に存在する個々の文字画像の総数  $Nt$  が所定の閾値  $Th3$  以下である場合 (S 2 7 で Y E S) にのみ、特別な文字出力形式の判定を行

うようにしている。

【0051】

まず、上記した注目単語についての文字出力形式の判定と同様に、注目文字画像ブロック内の個々の文字画像の中で、文字認識確度  $C$  が所定の閾値  $Th1$  より小さい文字画像データ候補の文字画像がいくつあるかをカウントし、 $Nb$  とする（S28）。

【0052】

注目文字画像ブロック内において、文字画像データ候補の文字画像の数  $Nb$  を文字画像の総数  $Nt$  で除した割合  $Nb / Nt$  が所定の閾値  $Th4$  より大きい場合（S29でYES）、注目文字画像ブロック内の全文字画像を切り出して文字画像データを作成する旨の文字出力形式の判定を行う（S30）。なお、閾値  $Th4$  は、例えば60%等、適宜設定することができ、例えば文字画像ブロック内の文字画像の総数  $Nt$  によって変化させることも可能である。さらに、割合  $Nb / Nt$  の閾値  $Th4$  を判定に使用する代わりに、文字画像データ候補の文字画像の数  $Nb$  を判定に使用することも勿論可能である。

【0053】

一方、文字画像ブロック内に存在する個々の文字画像の総数  $Nt$  が所定の閾値  $Th3$  より大きい場合（S27でYES）、上記した特別な文字出力形式の判定処理は行われない。

【0054】

つまり、本実施形態では、文字画像ブロックは、前述したように、例えば図6に示される矩形領域  $R1$ 、 $R2$  等のような、パラグラフ程度に文字画像が存在する集合をいう。ここで、閾値  $Th3$  は、適宜設定することができ、例えば若干少なく設定し、タイトルや脚注等の1～2行の文字画像からなる文字画像ブロックについて文字出力形式の判定処理を行うようにしてもよい。特に、タイトルには、様々なフォントが使われており、文字画像データと文字コードデータとが混在して出力されると、フォントの差異が目立つ可能性が高いので、文字画像ブロック単位で文字出力形式の判定を行う利点大きい。なお、あまり一般的でないフォントが使われていると文字認識確度も全体的に低下するので、原稿上のフォント

が文字コード認識の際に想定している一般的なフォントと異なる場合には、画像データ中の文字画像は、結果的に、文字コードデータに変換されずに文字画像データで出力されることになる。

#### 【0055】

そして、注目する文字画像ブロックの順番を示す変数  $k$  を繰り上げ (S31)、画像データ内の全文字画像ブロックについて、上記の処理を繰り返す (S32)。

#### 【0056】

このようにして、画像データ中の個々の文字画像について文字出力形式の判定処理が終了すると、図4に示すメインフローチャートに戻り、文字画像データの作成が行われる (S16)。すなわち、文字認識部31から出力される文字画像の位置情報にしたがって、画像データ中の文字画像を切り出して文字画像データを作成する。このとき、個々の文字画像ごとに外接矩形領域の画像データが切り出される。但し、単語や文字画像ブロック内の全文字画像を切り出して文字画像データを作成するような、連続した文字画像の集合を切り出す場合、例えば単語内の  $n$  番目の文字画像の位置情報、すなわち外接矩形の左上および右下の座標を  $(X1n, Y1n)$ 、 $(X2n, Y2n)$  とすると、 $(\text{Min}(X1n), \text{Min}(Y1n))$ 、 $(\text{Max}(X2n), \text{Max}(Y2n))$  で指定される単語内の全文字画像を含む矩形領域の画像データが切り出される。

#### 【0057】

ここで、文字画像データとして出力する文字画像を、原稿を読み取って得られた画像データとして元の位置に残さず、例えば単語単位で切り出すのは、文字コードデータと文字画像データとが重ならないようにするためである。つまり、一般に入力フォントと出力フォントとが全く同じである場合は少ないので、例えば文字画像から変換された文字コードデータが配置される位置と、入力された画像データ中の当該文字画像の位置とは、多少ずれてしまう。したがって、文字画像を、一文字画像ずつ、あるいは一単語ずつ切り出さずに画像データとして元の位置に残した場合、そのまま文字コードデータを配置すると、文字コードデータと原稿を読み取って得られた元の画像データ中の文字画像とが重なるといった問題

が生じ得る。そこで、本実施形態では、一文字画像ずつ、あるいは一単語ずつ切り出しておいて、文字画像データを作成し、例えば文字コードデータを出力フォントで配置した位置に続けて、切り出しておいた文字画像データを配置する。

#### 【 0 0 5 8 】

次に、文字消去処理が行われる（S 1 7）。すなわち、文字画像から文字コードデータまたは文字画像データを得た後、文字消去部 3 5 により、原稿を読み取って得られたカラーの画像データから、文字コードデータまたは文字画像データとされた後の元の文字画像が消去される。

#### 【 0 0 5 9 】

図 7 は、文字消去処理を説明するための図である。図 7 に示すように、個々の文字画像の外接矩形内における画像データ（図 7（A））から、文字画像に当たる部分を一旦消去し（図 7（B））、周囲の画像データで補完する（図 7（C））。消去する部分は、当該文字画像の外接矩形内におけるモノクロ 2 値画像データから黒を 2 ～ 3 画素膨張させる処理を行うことによって求める。なお、原稿を読み取って得られた元の画像データから文字画像に当たる部分を除去した後の画像データには、図 7（B）に示すように、文字画像以外の画像データ、すなわち、写真や線画に対応する画像、背景画像などが残っている。

#### 【 0 0 6 0 】

以上の処理を実施することにより、例えば図 8 に示すように、画像データ中の文字画像から、文字コードデータ D1 と文字画像データ D2 とが、また、非文字画像である図形画像をそのまま残した画像データ D3 が得られる。

#### 【 0 0 6 1 】

そして、これらの文字コードデータ、文字画像データ、および画像データは、所定のファイル形式でページ内に配置されて保存される（S 1 8）。なお、スキャン&プリントボタンが選択されている場合は、所定のデータがプリンタ 6 0 に送信されて用紙に印刷される。

#### 【 0 0 6 2 】

このように、本実施形態によれば、文字コードデータと文字画像データとの双方を、両者の位置ずれやフォントの差異を目立たなくしてユーザに与える違和感

を低減しつつ、自然に配置することが可能となる。

【0063】

図9は、本発明の第2実施形態に係る画像認識装置の文字出力形式判定処理を示すサブルーチンのフローチャートである。

【0064】

この第2実施形態は、上述した第1実施形態と画像認識処理における文字出力形式判定処理（図4のS15）の内容（図5参照）が相違している。以下、第2実施形態の画像認識処理について、第1実施形態との相違点を中心に説明する。

【0065】

スキャナ50で原稿を読み取ることによって得られた画像データに対し、図4に示すように、色空間の変換（S11）、OCR前処理（S12）、文字認識処理（S13）、および文字色検出処理（S14）が、上述した第1実施形態と同様に行われた後、画像データ中の個々の文字画像について文字出力形式の判定処理が行われる（S15）。すなわち、文字出力形式判定部33により、画像データ中の文字画像を、文字コードデータに変換するか、あるいは文字コードデータに変換せずに、文字画像を切り出して文字画像データを作成するかの判定が行われる。

【0066】

本実施形態の文字出力形式判定部33は、個々の文字画像ごとに、文字認識確度が所定値より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わるような特定文字画像が存在するか否かを検出し、以下に説明するように文字コードデータと文字画像データとを違和感なく出力させるための判定を行う。

【0067】

具体的には、図9に示すように、まず、注目する個々の文字画像の順番を示す変数nを初期化する（S41）。

【0068】

次に、注目文字画像の文字認識確度Cが所定の閾値Thより小さいか否かが判断される（S42）。文字出力形式判定部33は、あらかじめ、文字認識確度C

が所定の閾値  $T_h$  以上の文字画像を、文字コードデータに変換する候補の文字画像とし、文字認識確度  $C$  が所定の閾値  $T_h$  より小さい文字画像を、文字画像データを作成する候補の文字画像とする。そして、文字出力形式判定部 33 は、斜体の文字画像やカーニングが行われる文字画像に対する後述する特別な出力形式の判定処理が行われない場合には、個々の文字画像ごとに、候補とされた内容通りの文字出力形式の判定を行う。すなわち、文字認識確度  $C$  が閾値  $T_h$  以上の文字コードデータ候補の文字画像は、文字コードデータで出力し、文字認識確度  $C$  が閾値  $T_h$  より小さい文字画像データ候補の文字画像は、文字画像データで出力する旨の出力形式の判定が行われる (S43、S44)。なお、閾値  $T_{h1}$  は、適宜設定することができる。

## 【0069】

このように、文字出力形式判定部 33 は、基本的には、個々の文字画像ごとに、候補とされた内容通りの文字出力形式の判定を行うが、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わるような、斜体の文字画像やカーニングが行われる文字画像について、同様の判定を行うと問題が生じる。つまり、前述したように、用紙上に印刷された文字の一部において文字コードデータで印刷された文字と文字画像データで印刷された文字とが 2 重書きになったり、あるいは文字コードデータで印刷された文字の一部が文字画像データで印刷された文字によって上書きされて継ぎ目でずれてしまったりする。そこで、本実施形態では、個々の文字画像について、文字コードデータ候補の文字画像か、あるいは文字画像データ候補の文字画像かを決定した後に、以下に述べるように、さらに当該文字画像が斜体の文字画像やカーニングが行われる文字画像であるか否かを調べ、適切な文字出力形式の判定を行う。

## 【0070】

すなわち、注目文字画像の文字認識確度  $C$  が所定の閾値  $T_h$  より小さい場合、当該文字画像が斜体を呈しているか否かが判断される (S45)。文字画像が斜体を呈しているか否かの文字属性の情報は、文字認識部 31 で認識され、個々の文字画像ごとに文字属性情報として記憶されている。そして、当該文字画像が斜体を呈している場合 (S45で YES)、当該斜体の文字画像を含む文字画像集

合としての単語内の全文字画像を切り出して文字画像データを作成する旨の出力形式の判定を行う（S 4 6）。

#### 【0 0 7 1】

次に、注目文字画像が、1つ前の文字画像とカーニングが行われる組を形成しているか否かが判断される（S 4 7）。カーニングが行われる文字画像の組は、あらかじめデータとして記憶されており、例えば、W A、V A、W e 等が挙げられる。なお、カーニングが行われる組を形成しているとの判断には、組を成す個々の文字画像がそれぞれ外接矩形の位置情報を有していて、両者の隙間が詰められる処理がなされている場合だけではなく、組を成す2つの文字画像の隙間があらかじめ詰められていて、一体となった1つの外接矩形の位置情報を有している場合も含まれる。また、注目文字画像が、1つ後の文字画像あるいは前後の文字画像とカーニングが行われる組を形成しているか否かを判断するようにしてもよい。

#### 【0 0 7 2】

当該注目文字画像が1つ前の文字画像とカーニングが行われる組を形成している場合（S 4 7でY E S）、文字画像集合としてのカーニングが行われる組を成す両文字画像を切り出して文字画像データを作成する旨の出力形式の判定を行う（S 4 8）。なお、この場合（S 4 7でY E S）、カーニングが行われる組を含む単語内の全文字画像を切り出して文字画像データを作成する旨の出力形式の判定を行うことも可能である。

#### 【0 0 7 3】

そして、注目する文字画像の順番を示す変数 n を繰り上げ（S 4 9）、画像データ内の全文字画像について、上記の処理を繰り返し行う（S 5 0）。

#### 【0 0 7 4】

このようにして、画像データ中の個々の文字画像について文字出力形式の判定処理が終了すると、図4に示すメインフローチャートに戻り、文字画像データの作成が行われる（S 1 6）。すなわち、文字認識部 3 1 から出力される文字画像の位置情報にしたがって（図3参照）、画像データ中の文字画像を切り出して文字画像データを作成する。

## 【 0 0 7 5 】

このとき、個々の文字画像ごとに外接矩形領域の画像データが切り出される。但し、単語内の全文字画像や、カーニングが行われる組を成す両文字画像を切り出して文字画像データを作成するような、連続した文字画像の集合の切り出しは、以下のように行われる。

## 【 0 0 7 6 】

例えば、図 1 0 に示すように、文字画像が斜体を呈している場合、斜体の文字画像を含む単語内の  $n$  番目の文字画像の位置情報、すなわち外接矩形の左上および右下の座標を  $(X1n, Y1n)$ 、 $(X2n, Y2n)$  とすると、 $(Min(X1n), Min(Y1n))$ 、 $(Max(X2n), Max(Y2n))$  で指定される単語内の全文字画像を含む矩形領域の画像データが切り出される。

## 【 0 0 7 7 】

また、図 1 1 に示すように、文字画像がカーニングが行われる組を形成している場合、前の文字画像の位置情報、すなわち外接矩形の左上および右下の座標を  $(X1n, Y1n)$ 、 $(X2n, Y2n)$ 、後の文字画像の位置情報を  $(X1n+1, Y1n+1)$ 、 $(X2n+1, Y2n+1)$  とすると、 $(X1n, Y1n)$ 、 $(X2n+1, Y2n+1)$  で指定される矩形領域の画像データが切り出される。

## 【 0 0 7 8 】

ここで、文字画像データとして出力する文字画像を、原稿を読み取って得られた画像データとして元の位置に残さず、例えば単語単位で切り出すのは、前述したように、文字コードデータと文字画像データとが重ならないようにするためである。

## 【 0 0 7 9 】

次に、上記した第 1 実施形態と同様にして、文字消去処理 (S 1 7)、および所定の形式でのファイル作成 (S 1 8) が行われる。なお、スキャン&プリントボタンが選択されている場合は、所定のデータがプリンタ 6 0 に送信されて用紙に印刷される。

## 【 0 0 8 0 】

このように、本実施形態によれば、文字コードデータと文字画像データとの双



方を、両者が2重書きされたり継ぎ目でずれてしまったりすることを防止してユーザに与える違和感を低減しつつ、自然に配置することが可能となる。

## 【0081】

本発明は、上記した実施形態のみに限定されるものではなく、特許請求の範囲内において、種々改変することができる。

## 【0082】

例えば、上記実施形態では、入力された画像データの色空間を、RGB表色系からLab表色系に変換して画像処理するようにしたが、スキャナ50で読み取って得られた画像データのRGB表色系等、他の色空間を使用して画像処理を行うことも可能である。

## 【0083】

また、文字出力形式判定部33は、文字コードデータ変換候補の文字画像であるか否かを決定するための評価量として、文字認識確度を用いたが、この文字認識確度以外に、文字画像の位置情報や色情報から得られる文字らしさ（文字連続性）の度合いを組み合わせた評価量を用いることも可能である。

## 【0084】

また、注目単語あるいは文字画像ブロック内の全文字画像を切り出して文字画像データを作成する旨の判定を行う際に、例えば注目単語内における文字認識確度が所定値より小さい文字画像の数Nbを判定に使用したが、注目単語内における全文字画像の文字認識確度の平均値を判定に使用することも可能である。

## 【0085】

また、上記実施形態では、画像認識装置としてコンピュータを例に挙げて説明したが、本発明の画像認識装置はこれに限定されるものではなく、例えば同様の処理をスキャナで行ってコンピュータやプリンタ等に送信するシステム、スキャナから直接画像データを受信してプリンタで同様の処理を行うシステム、原稿の読み取りから用紙への印刷までの処理をカラーのデジタルコピー機ですべて行うシステム、等にも適用することが可能である。

## 【0086】

なお、上記した実施形態において、画像認識の制御は、上記した処理手順（図

4、図5等参照)を記述した所定のプログラムを主制御部11が実行することによって行われるものであり、この所定のプログラムは、コンピュータ読取可能な記録媒体(例えば、フロッピーディスクやCD-ROM等)によって提供されることもできる。また、この所定のプログラムは、上記各処理を実行するアプリケーションソフトウェアとして提供されてもよいし、コンピュータ、デジタルコピー機、スキャナ、プリンタ等の各情報機器や管理サーバの一機能として各情報機器や管理サーバのソフトウェアに組み込んでもよい。

【0087】

【発明の効果】

以上説明したように、本発明によれば、文字画像集合の中に、文字認識確度が所定確度より低い文字画像が含まれている場合、当該文字画像集合全体に対して文字コードデータへの変換を制止することによって、文字コードデータと文字画像データとが混在することによる両者の位置ずれやフォントの差異を目立たなくすることができ、ユーザに与える違和感を低減することが可能となる。

【0088】

また、文字認識確度が所定確度より小さく、かつ、文字画像の並び方向の端縁を通る並び方向に垂直な直線が隣接する文字画像と交わる特定文字画像が存在する場合、当該特定文字画像を含み個々の文字画像が複数集合して形成された文字画像集合全体に対して文字コードデータへの変換を制止することによって、文字コードデータと文字画像データとが2重書きされたり継ぎ目でずれてしまったりすることを防止することができ、ユーザに与える違和感を低減することが可能となる。

【図面の簡単な説明】

【図1】 本発明の第1実施形態に係る画像認識装置を含む情報機器システムの構成図である。

【図2】 図1に示されるコンピュータの概略構成を示すブロック図である。

【図3】 各文字画像の位置情報を説明するための図である。

【図4】 画像認識処理を示すフローチャートである。

【図 5】 図 4 に示される文字出力形式判定処理を示すサブルーチンのフローチャートである。

【図 6】 文字画像ブロックを説明するための図である。

【図 7】 文字消去処理を説明するための図である。

【図 8】 文字コードデータ、文字画像データ、および非文字画像である図形画像をそのまま残した画像データを配置した図である。

【図 9】 本発明の第 2 実施形態に係る画像認識装置の文字出力形式判定処理を示すサブルーチンのフローチャートである。

【図 1 0】 単語内の斜体の全文字画像を含む矩形領域の画像データが切り出された一例を示す図である。

【図 1 1】 カーニングが行われる組の両文字画像を含む矩形領域の画像データが切り出された一例を示す図である。

【図 1 2】 従来の画像認識装置により、文字コードデータと文字画像データとを配置した一例を示す図である。

【図 1 3】 従来の画像認識装置により、文字コードデータと文字画像データとを配置した一例を示す図である。

【符号の説明】

1 0 … コンピュータ（画像認識装置）、

3 1 … 文字認識部、

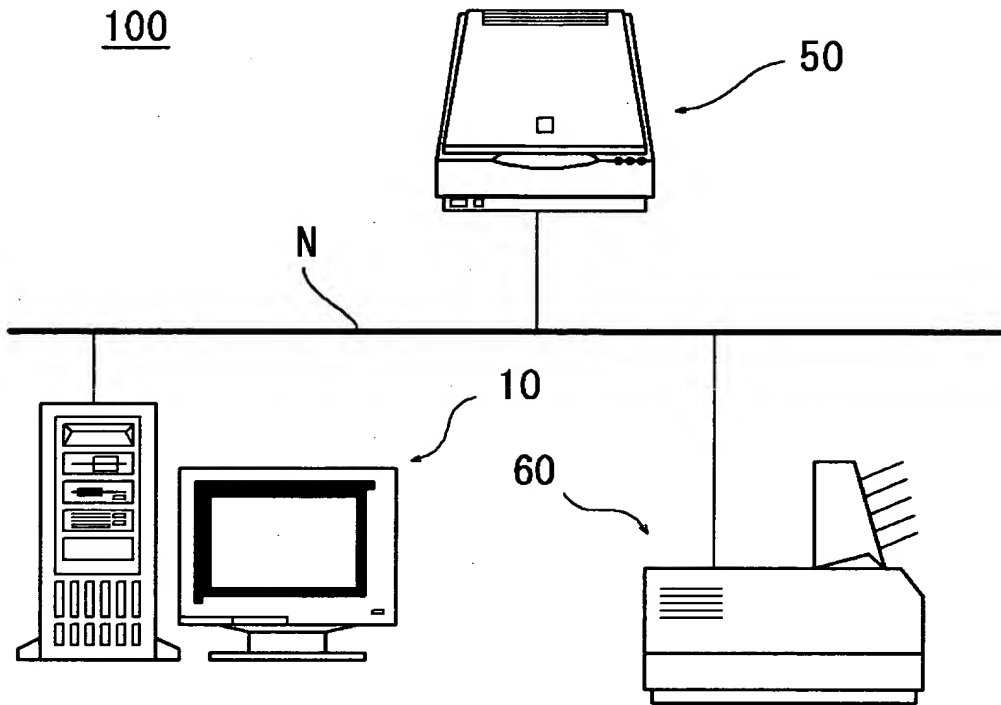
3 3 … 文字出力形式判定部、

3 6 … ファイル作成部。

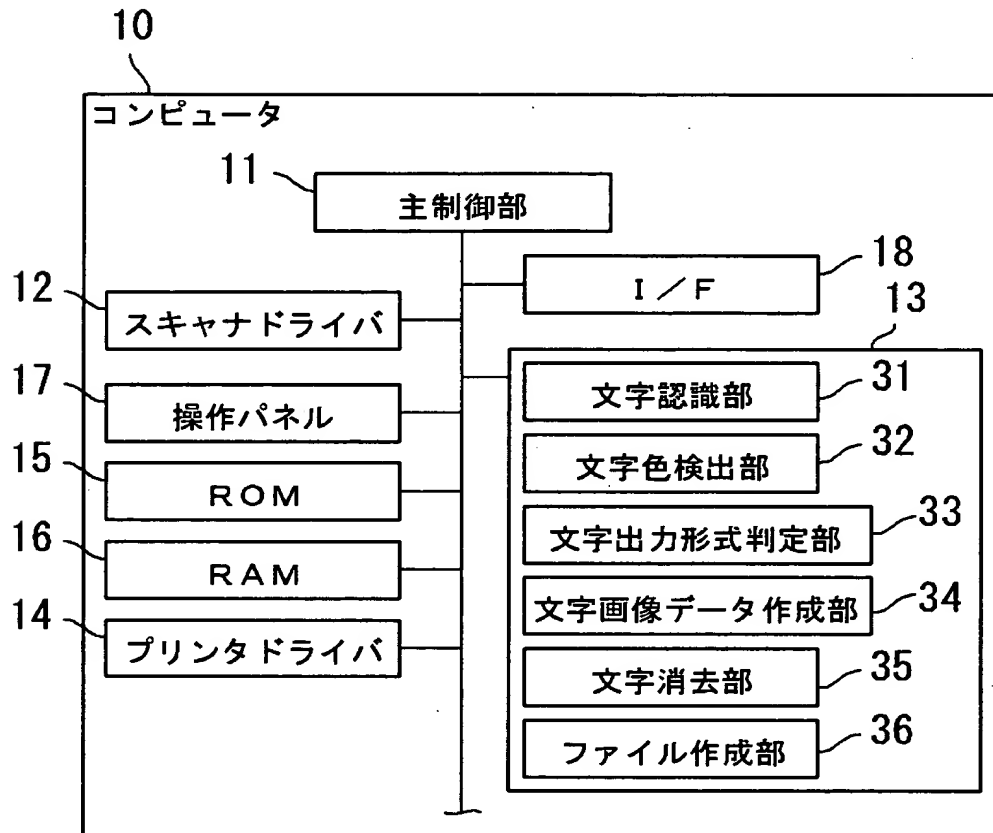
【書類名】

図面

【図 1】

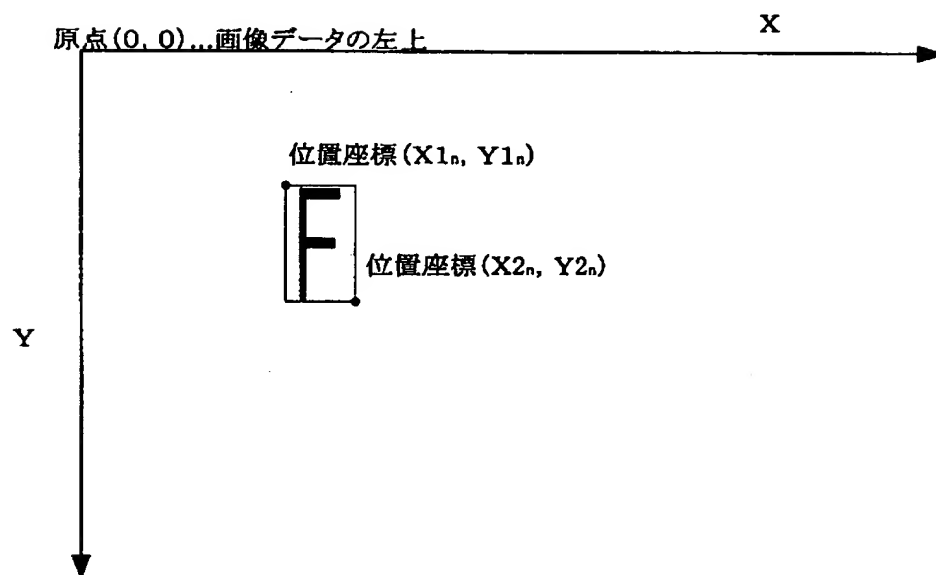


【図 2】

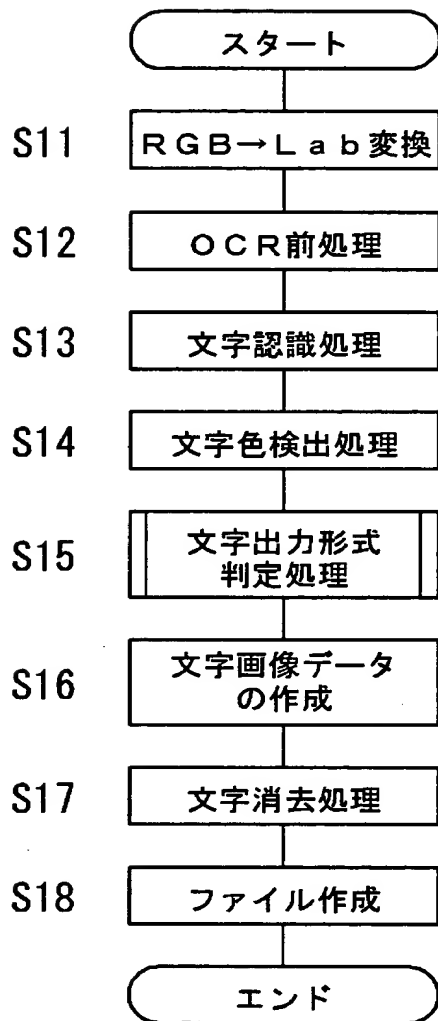


【図 3】

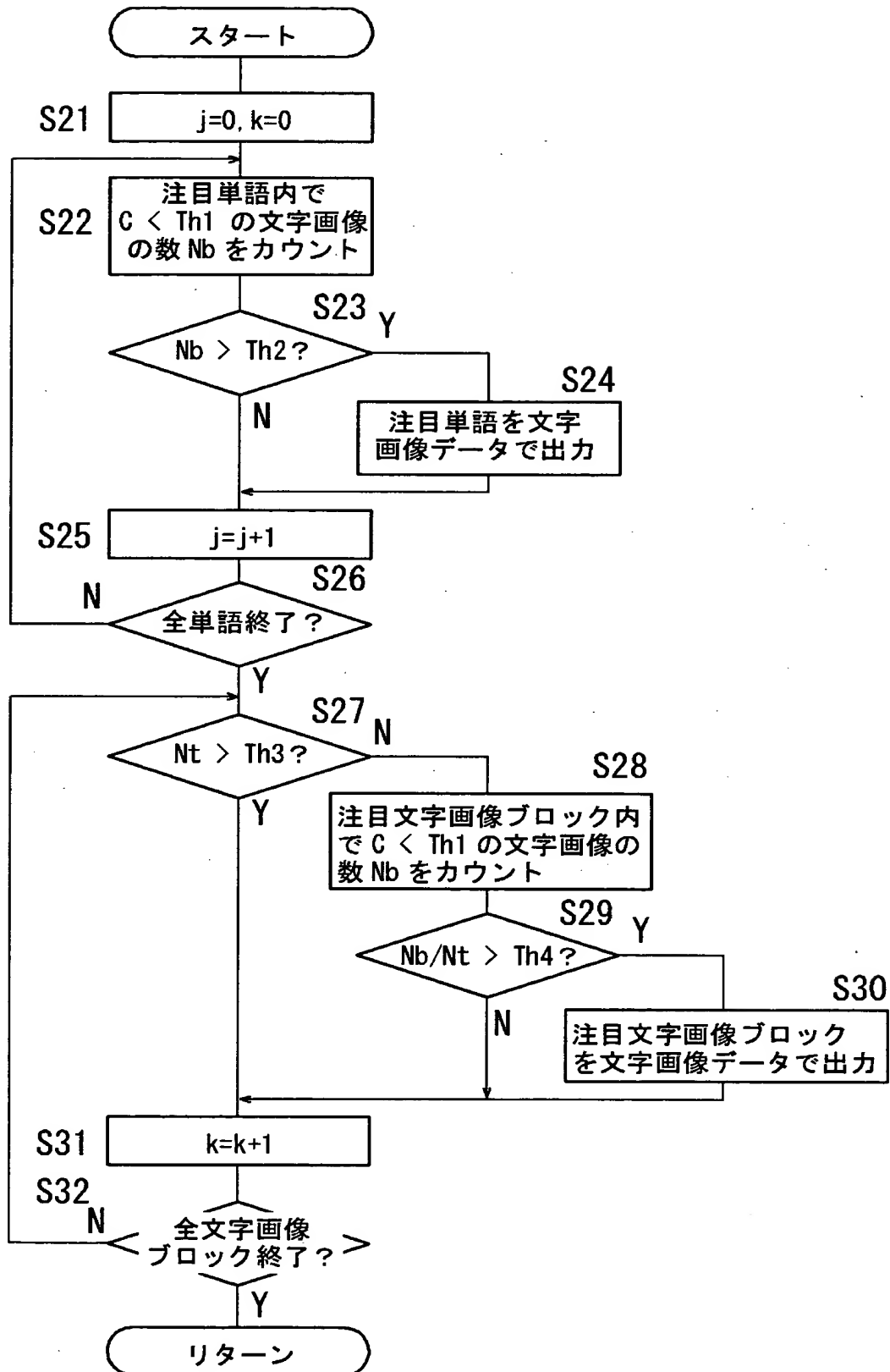
Font type



【図 4】



【図 5】





【図 6】

R1

**Special Offer**

R2

Speed up productivity and decrease time-to-market by  
calling today for your Professional Image Processing Kit  
and this incredible special offer.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

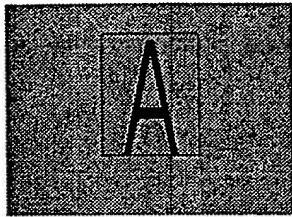
\_\_\_\_\_

\_\_\_\_\_

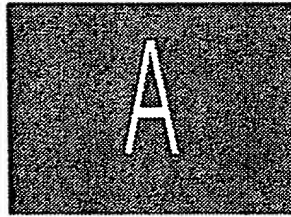
\_\_\_\_\_

【図7】

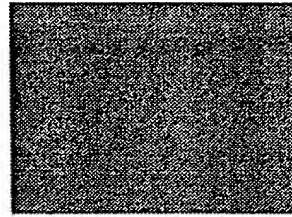
(A)



(B)



(C)



【図8】

D2

D1

D3

**It's Perfect!**

**Maxy's** fresh-picked fruit is always the perfect gift.

Everyone will appreciate your thoughtfulness.

Pineapples	#001-ca	\$30.00
Navel Oranges	#002-cc	\$25.00
Red Grapes	#003-cc	\$20.00

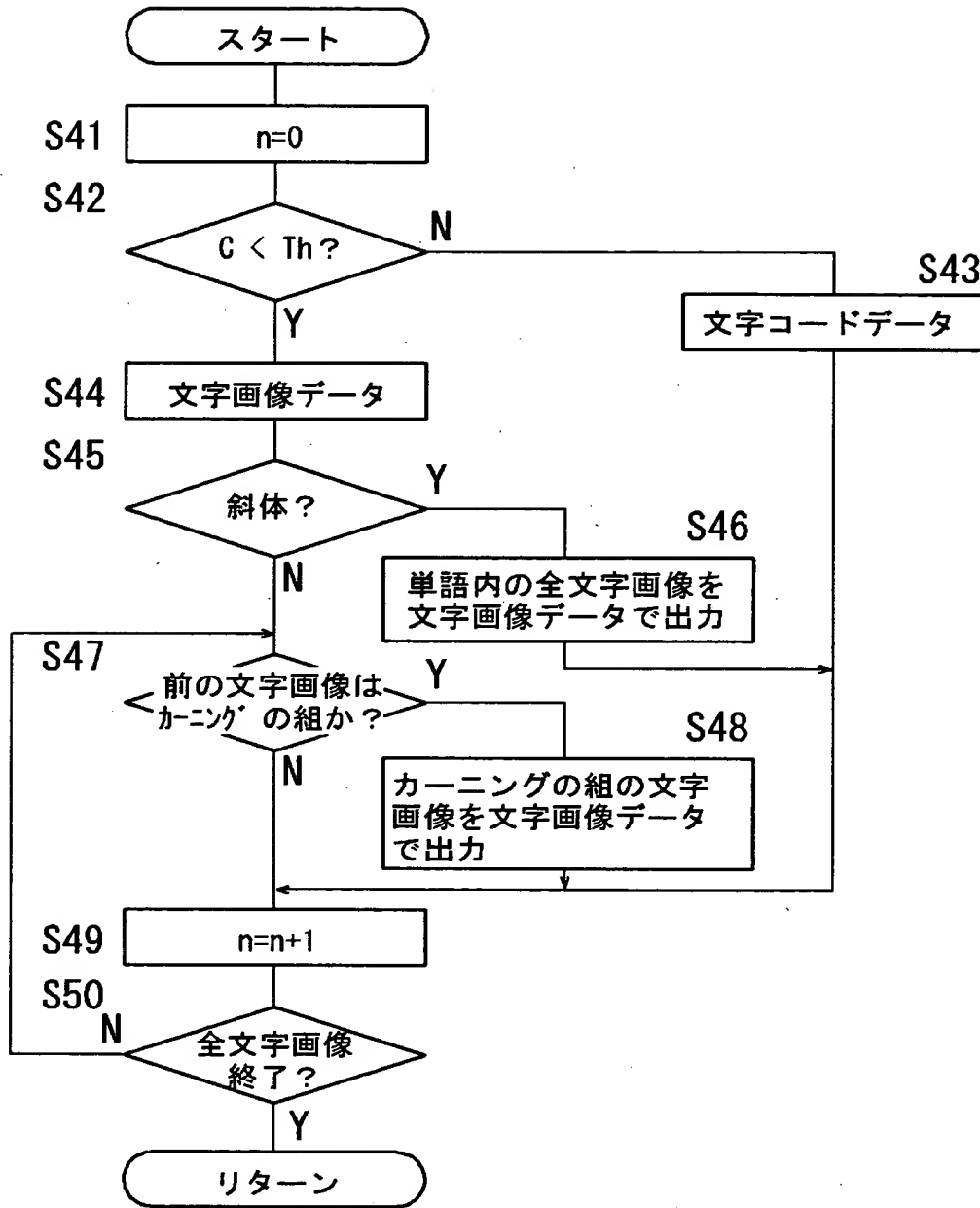
Order now! Your satisfaction is guaranteed.

Phone: (800)123-4567

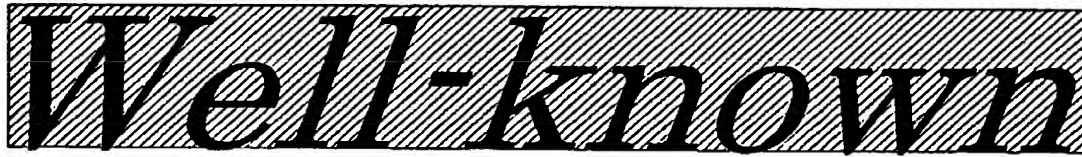
Internet: <http://www.maxy-fruit.dot.com>

D2

【図 9】



【図 1 0】



【図 1 1】

位置座標 (X1n, Y1n)



位置座標 (X2n+1, Y2n+1)

【図 1 2】

**Special Offer**

Speed up productivity and decrease time-to-market by  
calling today for your Professional Image Processing Kit  
and this incredible special offer.

【図 1 3】

(A)

Well-known

(B)

Well-known

【書類名】                      要約書

【要約】

【課題】    画像データ中の文字画像から文字コードを認識して得られた文字コードデータと文字画像から文字コードデータに変換することなく切り出された文字画像データとが混在する画像を出力する画像認識装置において、双方の混在によりユーザに与える違和感を低減する。

【解決手段】    文字出力形式判定部 3 3 は、例えば単語等の文字画像集合の中に、文字認識確度が所定値以上の文字コードデータ候補の文字画像と、文字認識確度が所定値より小さい文字画像データ候補の文字画像とが混在する場合、当該単語内の全文字画像を文字コードデータに変換することなく切り出して文字画像データを作成する旨の判定を行う。

【選択図】                      図 2

出 願 人 履 歴 情 報

識別番号 [000006079]

1. 変更年月日 1994年 7月20日

[変更理由] 名称変更

住 所 大阪府大阪市中央区安土町二丁目3番13号 大阪国際ビル  
氏 名 ミノルタ株式会社